1

2

3

4

5

6

7

8

9

10

UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA
SAN FRANCISCO DIVISION

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

IN RE: FACEBOOK, INC. CONSUMER
PRIVACY USER PROFILE LITIGATION,

This document relates to:

ALL ACTIONS

MDL NO. 2843

CASE NO. 3:18-MD-02843-VC-JSC

HON. VINCE CHHABRIA
HON. JACQUELINE SCOTT CORLEY
COURTROOM 4 – 17TH FLOOR
SPECIAL MASTER, DANIEL GARRIE, ESQ.

ORDER FOLLOWING MARCH 9, 2022
HEARING REGARDING PLAINTIFFS'
MOTION TO COMPEL PRODUCTION
OF PLAINTIFF DATA

ORDER FOLLOWING MARCH 9, 2022 HEARING REGARDING PLAINTIFFS' MOTION
TO COMPEL PRODUCTION OF PLAINTIFF DATA

1

2   No later than April 5, 2022, Facebook is to submit a proposed protocol for producing named

3   Plaintiff user data pursuant to Judge Corley's Discovery Order No. 9 (Dkt. 557) beyond what has been

4   produced to date. No later than April 11, 2022, Plaintiffs are to submit a response to Facebook's

5   proposed protocol. Facebook's proposal should include, if appropriate, the following types of

6   information.

7   • Data flow diagrams of the systems from which Facebook searches and produced Named Plaintiff

8   Data.

9   • Functional descriptions and interdependencies for DataSwarm tasks that process Named Plaintiff

10   Data.

11 • Descriptions of the schemas, tables, columns, and data types for Named Plaintiff Data that is

12   produced.

13   Facebook's proposal should identify any Facebook systems and the Named Plaintiff Data that it

14 will not produce from and include an explanation as to why it will not produce from such systems (i.e.,

15 burden, costs, duplicative, etc.)

16 The Special Master Garrie suggests that Facebook's protocol include producing the Named

17 Plaintiff Data from the following.

18

19 • **The Associations Objects (TAO)** – Facebook is to identify and produce Named Plaintiff Data
      stored in TAO that is not present in the Named Plaintiff's DYI file and it should include, if
20   applicable, the following: (a) the Objects and Associations in TAO associated with the Named
      Plaintiffs, but are either partially or not included in the DYI file (i.e. Named Plaintiff Data that is
21   not exported in DYI file but stored in TAO). For example, data about a Page (as defined by
      Facebook) that the Named Plaintiff interacted with is stored in a MySQL table in TAO called Y,
22   which was not included in the DYI file. Facebook identifies table Y and produces the
      rows/columns from table Y with the Named Plaintiff Data relating to the page.
23

24 • **Hive** -- Facebook is to query the Hive to identify the tables that store Named Plaintiff Data using
      the identifiers including the following: User ID (UID), Replacement ID (RID) Separable ID
25   (SID), App-Scoped ID (ASID). For each table identified Facebook is to search the tables for the
      associated Named Plaintiff Data. (i.e., tables mapping user identifiers to ad segmentation data
26   such as US Political spectrum segments). For example, a Facebook engineer writes a process that
      stores user data, including Named Plaintiff Data, in the Hive in table X and the table has a
27

28

**ORDER FOLLOWING MARCH 9, 2022 HEARING REGARDING PLAINTIFFS' MOTION
TO COMPEL PRODUCTION OF PLAINTIFF DATA**

column "RID." Facebook would produce the Named Plaintiff Data that is stored in table X. This effort should exclude any of the analysis done in relation to DataSwarm below.

- **DataSwarm** -- Facebook is to query DataSwarm Tasks to identify Task Definitions which involve Named Plaintiffs Data, using known identifiers such as User ID (UID), Replacement ID (RID) Separable ID (SID), App-Scoped ID (ASID). Facebook will review each of those Task Definitions and then search the sources/destinations identified in the task for the Named Plaintiff Data. See Special Master Hearing Transcript 3/9 p.35 5-13. For example, Facebook queries the DataSwarm Tasks and identifies a Task that uses the UserID to pipe data to Laser and stores that data in table Y. Facebook is to search and produce from table Y all Named Plaintiff Data (columns/row/schemas) that was not included in the DYI file.

Written Questions

Facebook is also to submit answers to the following questions and requests for documentation to the Special Master on or before April 1, 2022.

<u>Hive</u>

- How does ad impression and ad click data for Facebook users get into Hive?

- What tables store ad impression and ad click data for Facebook users?

- What data pipelines[1] are used to analyze ad impression and ad click data for Facebook users? Where is the final output of these data pipelines stored (e.g. Hive table names, TAO, etc.)?

- Identify a list of Hive tables containing columns that store a UID, RID, SID, ASID, or other means of identifying a Facebook user. The list is to include tables that were active during the relevant time period.

- What is the estimated time and cost to produce data for the Named Plaintiffs from Hive?

<u>Ads Interests</u>

- How does Facebook determine ads interests for a user based on what the user views (i.e. are ad interests based on what the user views on Facebook or other Internet activity)?

- Does Facebook track user activity across the Internet using cookies? If so, what cookies does Facebook use? Provide a statement explaining the use of cookies in tracking user activity on and

---

[1] Data pipeline as used here refers to an end to end data process using data tasks in a dependency chain.

ORDER FOLLOWING MARCH 9, 2022 HEARING REGARDING PLAINTIFFS' MOTION TO COMPEL PRODUCTION OF PLAINTIFF DATA

off the Facebook platform to create behavioral data about users. Provide documentation on the use of _fbp, _fbc, and DATR cookies. Explain whether the scope of tracking user activity includes on or off platform activity, or both.

- How are ads interests associated with a particular user? Where is ads interests data for individual users stored? Facebook is to describe whether ad interests data can be associated with a specific user via UID, RID, SID, ASID, or other means and whether it is included in the DYI file.

- What is the estimated time and cost to produce ad interests data for the Named Plaintiffs?

<u>Contracts</u>

- Is the data referenced in the contracts with Netflix, Microsoft, and/or YouTube that Facebook provided to Special Master Garrie for *in camera* review in connection with the Named Plaintiff Data hearings included in the DYI file? If not, Facebook is to specify what data is not included and where such data stored.

Scenarios

No later than April 4, 2022, Facebook is to submit documentation sufficient to describe the data collected both on and off platform or provided by Third Parties in the following scenarios and provide written responses to the questions below.

- Exhibit A to Plaintiff's Questions re: Data Collection and Use indicates that Facebook used predictive algorithms to generate five political segments for Facebook users (Very Liberal, Liberal, Moderate, Conservative, and Very Conservative) based on demographic, psychographic, and behavioral signals from Facebook user data.

   o What are the inputs into these algorithms (i.e. what are the demographic, psychographic, and behavioral signals used to generate the political segments)? Are these inputs provided by users or derived by Facebook?

   o How are the psychographic signals computed (e.g. how is the psychographic signal "High Dollar Religious Donor" determined)?

ORDER FOLLOWING MARCH 9, 2022 HEARING REGARDING PLAINTIFFS' MOTION TO COMPEL PRODUCTION OF PLAINTIFF DATA

- o Is information regarding identifiable ethnic affinities provided by users or derived by Facebook? How is ethnic affinity derived?

- o Where is political segmentation data for Facebook users stored?

- o Is political segmentation determined for a Facebook user as part of a data process that runs on a regularly scheduled basis or evaluated in real time when an ad is served?

- o Is political segmentation associated with a Facebook user if possible (i.e. via UID, RID, SID, ASID, or another identifier that can be mapped to a user)? If so, explain how the political segmentation is associated with a Facebook user.

- o Is an individual's assigned political segment part of the DYI file?

IT IS SO ORDERED.

Tuesday, March 22, 2022

_____

Daniel Garrie
Discovery Special Master

ORDER FOLLOWING MARCH 9, 2022 HEARING REGARDING PLAINTIFFS' MOTION TO COMPEL PRODUCTION OF PLAINTIFF DATA